Working Paper Series

Max-Sebastian Dovì, Gerrit Koester,
Christiane Nickel

Addressing the endogeneity of slack
in Phillips Curves

Abstract

Endogeneity of the labour market slack in reduced-form Phillips Curves (PCs) is usually addressed either by including proxies for omitted supply shocks, or by using instrumental variables. Using the Kiviet (2020) Kinky Least Squares estimator, we find evidence that supply-shock proxies should not be omitted from PCs, and that many popular instrumental variables seem to be invalid. We estimate a standard backward-looking wage Phillips Curve by Kinky Least Squares and find that unless a large *negative* correlation between the slack variable and the error term is assumed, the coefficient of the slack variable is significantly negative.

Keywords— Phillips Curves; instrument-free inference; limited-information inference
JEL classification— C1, E3

Non-Technical Summary

The relationship between economic slack and inflation as conceptualised in the Phillips Curve is of core importance for the working of monetary policy. The wage Phillips Curve relationship implies that decreases in economic slack would put upward pressure on wages, while increases in economic slack should lead to downward pressure, which should be reflected in a negative coefficient between slack and wages. Such Phillips-Curve relationships are often assessed based on reduced-form estimations (e.g., Nickel et al. (2019)).

Estimating (reduced-form) Phillips Curves faces the problem of endogeneity of the slack variable. This endogeneity of the slack variable could result from omitted variables such as supply shocks that are correlated with the slack variable, or for example monetary policy (partially) neutralising the effects of demand shocks on prices and wages. It is usually addressed either by including proxies for omitted supply shocks, or by using instrumental variables.

We argue that approaches that augment standard Phillips Curves with supply shocks require the unverifiable assumption that all the endogeneity is due to the variation in the omitted supply-shock proxies, while IV-based methods require making questionable assumptions on the uncorrelatedness of the chosen instrumental variables with the error term. We also show that both approaches suffer from specification uncertainty, that cannot be addressed by standard 'thick modelling' approaches, since it is unclear how to aggregate the inferential results from the very many different specifications available.

The main methodological contribution of this paper to the literature on Philips Curves estimation consists in advocating the use of the Kiviet (2020) Kinky Least Squares estimator, which uses prior information the researcher may informally hold on the correlation between the variable they want to conduct inference on and the error term of the model. This prior information can be based on knowledge the researcher has on omitted variables, or unmodelled economic mechanisms that lead to the variable being simultaneously determined with the error term. By restricting attention to economically plausible parts of the parameter space of the correlation coefficient between the endogenous variable and the error term, it is possible to conduct endogeneity-robust *and* instrument-variable-free inference. Hence, rather than trying to explicitly account for the endogeneity present in the model, Kinky Least Squares provides bounds on how 'extreme' the endogeneity present in the model has to be before central qualitative conclusions from a regression analysis are overturned. Furthermore, Kinky Least Squares estimation makes it possible to directly test whether candidate supply-shock proxies or candidate instrumental variables have been incorrectly omitted from the specification using standard t-tests and Wald tests.

Applying Kinky Least Squares to a standard backward-looking wage Phillips Curve, we find evidence that supply-shock proxies should not be omitted from Phillips Curves, and that many popular instrumental variables seem to be invalid. Estimating a standard backward-looking wage PC for the euro area by KLS, we find that unless a large negative correlation between the slack variable and the error term is assumed, the coefficient of the slack variable is indeed significantly negative. Since the correlation coefficient of the slack variable and the error term is often thought to be *positive* (due, for instance, to omitted supply shocks, which tend to have especially large effects in times of large economic slack as argued in Gordon (2011), or monetary policy, which becomes more active with increasing levels of economic slack as argued in Eser et al. (2020)), we interpret this result as evidence supporting the view that the wage Phillips Curve in the Euro Area is alive and well.

1   Introduction

Due to its role in the transmission of monetary policy, the causal effect of economic slack on wage or price inflation is perhaps one of the most widely debated questions in academic and applied macroeconomics. The vast literature on the topic reports a wide range of estimates for the effect of slack on wages and prices, which remains a central question in macroeconomics.

One important reason for the wide range of estimates in the literature could be the likely endogeneity of the slack variable (i.e. the non-zero correlation of the slack with the error term), and the steps taken to address it. Indeed, if the slack variable were exogenous and the error term in the Phillips Curve (PC) well-behaved, the coefficient on the slack variable could be consistently estimated by reduced-form OLS, and the results reported in the literature (in the absence of structural breaks) should differ only modestly.

The likely endogeneity of the slack variable could result from omitted variables that are correlated with the slack variable (such as supply shocks, see Gordon (2011)). Monetary policy could also play an important role, since any (successful) policy action designed to offset demand shocks would lead to a co-movement of slack and the error terms in a reduced-form PC (see, e.g., McLeay and Tenreyro (2019)). Such factors could lead to fundamental differences in estimates from reduced-form approaches and (semi-)structural models (see, e.g., Eser et al. (2020)).

Broadly speaking, the literature on reduced-form estimation of PCs has addressed the likely endogeneity of the slack variable in two ways.[1] The first approach is based on adding to the PC those variables that could affect the transmission of slack to wage growth. In the literature, this has predominantly taken the form of adding variables that ostensibly proxy supply shocks, such as commodity inflation, monetary policy, or an economy's exchange rate (Mazumder, 2014; Gordon, 2011). The second approach involves using instrumental variable (IV) methods, by exploiting some type of exclusion restriction. This approach gained popularity in the wake of the research conducted on structural PCs, but is also used in so-called 'semi-structural' and reduced-form estimation (see, e.g. Kleibergen and Mavroeidis (2009), Dufour, Khalaf, and Kichian (2006), Bulligan and Viviano (2017), and Mazumder (2014)).

Even if these methods could fully address the endogeneity of the slack variable, the high dimensionality of macroeconomic data implies that both approaches are still inherently arbitrary. This is because there is substantial 'specification uncertainty' (Mavroeidis, Plagborg-Møller, and Stock, 2014), whereby there is no *a priori* reason for preferring one particular set of available supply-shock proxies or IVs over another.

Against this background, this paper proposes an alternative approach to address the endogeneity of slack in a reduced-form wage PC for the euro area (EA).

The first contribution of this paper is to illustrate the problem of specification uncertainty of a standard backward-looking reduced-form wage PC for the EA. We consider a wage PC, as it has received scant attention (relative to price PCs) in the past literature, while noting that the concerns related to supply-shock proxies and questionable IVs raised in the literature on price PCs extend to the case of wage PCs. Considering widely-used supply-shock proxies and IVs, we show that inference on the slack coefficient depends on which of these are chosen. Since it is not possible

---

[1]Due to the focus of this paper on direct (parametric) estimation of PCs, we do not consider recent 'indirect' approaches to establishing relations between labour market slack and inflation through VARs and DSGE models (see, e.g., Del Negro et al. (2020)).

to consider all supply-shock proxies or IVs and their transformations, this makes inference on the coefficient of the slack variable dependent on an arbitrary choice of supply-shock proxies or IVs.[2]

As a second contribution, we show how the recently proposed 'Kinky Least Squares' (KLS) estimator of Kiviet (2020) can be used for specification testing in the context of PCs. KLS makes it possible to directly test for the correlation of the error term with any variable by bias-adjusting the OLS estimator for a given degree of endogeneity. Consequently, we can formally test whether a given supply-shock proxy has been incorrectly omitted from a given PC.[3] Equivalently, this test can also be treated as a test of IV validity, so that for the first time in the vast literature on PCs, we are able to directly test the exclusion restrictions of IVs.[4] On either interpretation, this test is especially useful in reduced-form modelling of PCs, since by definition, there is no theoretical or structural reason for considering one set of supply-shock proxies over another or one set of IVs over another.[5]

As a third contribution, we propose to use the KLS estimator also to conduct endogeneity-robust inference on the coefficient of the slack variable without using any supply-shock proxies or IVs. Rather than relying on ad-hoc and not very transparent arguments for preferring one set of supply-shock proxies or IVs over another, we propose to identify the likely sign and magnitude of the coefficient of the slack variable by leveraging the widely-held intuition that the correlation coefficient of the slack variable and the error term is likely *positive* (due, for instance, to omitted supply shocks, which tend to have especially large effects in times of large economic slack as argued in Gordon (2011), or monetary policy, which becomes more active with increasing levels of economic slack as argued in Eser et al. (2020)). Thus, rather than trying to explicitly account for the endogeneity present in the PC, we provide bounds on how 'extreme' the endogeneity has to be before the central qualitative conclusions drawn from a standard OLS analysis are overturned. We find that for non-negative as well as moderately negative values of the correlation coefficient between the slack variable and the error term, the coefficient on the slack variable is significantly negative, supporting the view that the wage PC in the EA is alive and well.

Although the focus of this paper is on conducting inference on the coefficient of the slack variable in wage PCs, the main issues raised and solutions proposed in this paper are relevant to many other applications in macroeconomics, such as inference on Taylor rules, and Euler equations. This paper is hence closely related to recent work by Carvalho, Nechio, and Tristao (2019), who use a structural model to argue that Taylor rules should be estimated by OLS, since any bias caused by endogeneity is likely to be small. Our paper extends their line of argument in two important ways. First, if, as in Carvalho, Nechio, and Tristao (2019), a structural model is at hand to guide the researcher, KLS makes it possible to exploit both the sign and magnitude of the correlation coefficients between the endogenous variables and the error term for inference. In contrast, the OLS

---

[2]It should be noted that 'thick modelling' approaches (that consider very many different, if not all, specifications as in Koester et al. (2021) and Ciccarelli and Osbat (2017)) do not provide a satisfying solution to this problem in the given context, since they do not account for estimation uncertainty (beyond Bonferonni corrections that are bound to be powerless given the large number of possible specifications). It should also be noted that recently developed 'big data' econometric methods such as Dovì (2021) and Chernozhukov et al. (2019) are not able to qualitatively solve the issue of specification uncertainty. This is because even these methods are unable to handle an arbitrarily large number of variables (see the discussion in Dovì (2021)).

[3]Previous OLS-based testing of whether a set of supply-shock proxies are significant in a given PC can only provide suggestive evidence, since the validity of such a test hinges on all variables being exogenous, including the slack.

[4]Previously, any test of IV validity was limited to tests of overidentifying restrictions, see Kleibergen and Mavroeidis (2009) and Mavroeidis, Plagborg-Møller, and Stock (2014).

[5]In inference on structural PCs this issue is not as pronounced, since the PC is correctly specified, and any exclusion restriction is valid by microfoundational assumption.

approach advocated by Carvalho, Nechio, and Tristao (2019) makes no formal use of the sign of the correlation between the endogenous variables and the error term. Second, our approach remains valid even in the absence of a structural model. All that is required is an informal (non-parametric) 'prior' on the degree of endogeneity. The remaining assumptions are all testable. This recommends our approach in reduced-form (policy) settings, where practitioners may have intuition grounded in qualitative economic theory, without being committed to a fully-specified macroeconomic structure.

This paper is organised as follows. Section 2 specifies the PC used throughout this paper. Section 3 introduces the methodology and the data. Section 4 presents the results. Section 5 concludes.

2 Model

In this paper we consider the following reduced-form PC:

$$\pi_t^w = \lambda u_t + \sum_{j=1}^{q_w} \gamma_{wj} \pi_{t-j}^w + \sum_{j=1}^{q_p} \gamma_{pj} \pi_{t-j}^p + \epsilon_t, \tag{1}$$

where u_t is the labour market slack variable with coefficient λ, π_t^w and π_t^p are the wage and price inflation, γ_{wj} and γ_{pj} are the coefficients on the jth lag of the wage and price inflation, respectively, and ϵ_t is a mean-zero error term (to simplify notation, we treat the variables as demeaned throughout). This reduced-form wage PC with backward-looking inflation expectations is also the workhorse model, e.g. in Nickel et al. (2019), for the analysis of wage developments in the euro area.

We focus on reduced-form PCs out of a desire for generality. Indeed, inference on PCs embedded in a structural model can be seen as a special case, where primitive assumptions imply that the PC is correctly specified, the error terms are well-behaved, the IVs are valid, and the degree of endogeneity is determined by the model's 'deep' parameters. We omit forward-looking variables out of simplicity, and without compromising our ability to conduct inference on the coefficient of the slack variable. Indeed, since under certain (testable) assumptions the KLS-based approach we propose addresses the likely endogeneity of the slack variable, the results will be robust to potentially omitted variables. These could include, for example, forward-looking inflation expectations.[6]

3 Methodology and Data

For all approaches, it is useful to cast the PC specified in Equation (1) in matrix form:

$$\begin{aligned} \pi &= \lambda u + X\xi_1 + \epsilon \\ &= X^+\beta + \epsilon, \end{aligned} \tag{2}$$

where π is the $T \times 1$ stacked vector with element π_t^w, u is the $T \times 1$ stacked vector with element u_t, X is the $T \times (q_w + q_p)$ matrix with columns corresponding to the $T \times 1$ stacked vectors of $\pi_{t-j_w}^w$, $\pi_{t-j_p}^p$ for $j_w = 1, \ldots, q_w$, and $j_p = 1, \ldots, q_p$, respectively, ξ_1 is the $(q_w + q_p) \times 1$ corresponding vector of coefficients, ϵ is the $T \times 1$ stacked vector with element ϵ_t, $X^+ = [u\ X]$, and $\beta = [\lambda, \xi_1']'$.

---

[6]Nickel et al. (2019) find that backward-looking expectations seem to be more relevant to explain wage growth in the euro area through the lens of the Philips Curve than forward-looking expectations.

Throughout, we assume that ϵ_t is independent across t. Under this assumption, we can treat the matrix of lagged (predetermined) variables, X, as exogenous. However, u is allowed to be correlated with ϵ.

3.1 Single-Equation Estimation

The first approach considered in this paper is to directly proxy for supply shocks in the PC, i.e. estimate the model given by

$$\begin{aligned}
\pi &= X^+\beta + S\delta + \tilde{\epsilon} \\
&= W\zeta + \tilde{\epsilon},
\end{aligned} \tag{3}$$

where S is a $T \times (q_s + 1)g$ matrix consisting of the $g \times 1$ vector of contemporaneous supply-shock proxies, q_s the number of lags, δ the $(q_s + 1) \times 1$ corresponding coefficient vector, $W = [X^+\ S]$, and $\zeta = [\beta', \delta']'$. The crucial identifying assumption is that by including these additional supply-shock proxies, all the variables on the RHS are uncorrelated with the new error term, $\tilde{\epsilon}_t$, i.e.

$$\mathbb{E}\left[W'\tilde{\epsilon}\right] = 0.$$

Assuming that other standard regularity conditions hold and that $q_s g$ is fixed (i.e. $q_s g << T$), this moment condition makes it possible to conduct inference on Equation (3) by OLS. Such OLS-based approaches first gained popularity with Gordon (1982), and are sometimes termed 'triangle models', since they include demand variables (the slack variable), inertia (lagged inflation variables) and supply variables (the supply-shock variables). They remain popular both in academic settings (Ball and Mazumder, 2019; Mazumder, 2014; Gordon, 2011) and policy settings (Cunningham, Rai, and Hess, 2019).

3.2 Instrumental Variable Estimation

The second popular approach to address the endogeneity of the slack variable is to use IVs. This involves exploiting moment conditions given by

$$\mathbb{E}[Z'\epsilon] = 0,$$

where Z is a $T \times k$ matrix with columns corresponding to (any) predetermined variable not included in Equation (2).

A well-known problem in the context of PCs is that the IVs are likely to be weak, i.e. they are able to explain only a small portion of the exogenous variation in the endogenous variable (Mavroeidis, 2005; Ma, 2002). Therefore, to avoid biases caused by pre-testing, we consider the Anderson and Rubin (1949) (AR) statistic, which is completely robust to (arbitrarily) weak identification, and has other attractive properties vis-à-vis other more sophisticated weak-identification robust methods (see Dufour (2003) for a discussion).

Confidence sets are constructed by inverting the AR statistic under a given null hypothesis. In our application, we have a single endogenous variable, so that the relevant null hypothesis is given by

$$H_0 : \lambda = \lambda_0 \text{ vs. } H_1 : \lambda \neq \lambda_0.$$

Constructing $(1-\alpha)$ confidence sets by inverting the test statistics involves specifying a (sufficiently fine) grid of values of λ_0, and evaluating the test statistics at each of these values. The hypotheses that cannot be rejected at the α level are included in the $(1-\alpha)$ confidence set, whereas the hypotheses rejected at the α level are not.

The AR statistic is given by

$$AR = \frac{\frac{1}{k}\epsilon_0^{*\prime}P_{Z^*}\epsilon_0^*}{\frac{1}{T-k}\epsilon_0^{*\prime}M_{Z^*}\epsilon_0^*},$$

where $\pi^* = M_X\pi$, $u^* = M_X u$, $Z^* = M_X Z$, $\epsilon_0^* = \pi^* - \lambda_0 u^*$, $M_X = I - X(X'X)^{-1}X'$ $P_{Z^*} = Z^*(Z^{*\prime}Z^*)^{-1}Z^*$, and $M_{Z^*} = I - P_{Z^*}$. Assuming that k is fixed (i.e. $k << T$), and that other regularity conditions hold (Dufour, Khalaf, and Kichian, 2006, Appendix B), the AR statistic converges in distribution to $1/k$ times a χ_k^2 distributed random variable.

3.3   Kinky Least Squares Estimation

The third approach considered in this paper is KLS. The crucial quantity in KLS is the vector of correlations between the columns in X^+, and the error term ϵ. This vector of correlations will be denoted by the $(1 + q_w + q_p) \times 1$ vector $\rho_{x+\epsilon}$. Letting r_j denote the assumed (potentially non-zero) correlation coefficient between X_j^+ and the error term ϵ, the assumed correlation coefficients vector will be given by the $(1 + q_w + q_p) \times 1$ vector $r = [r_1, 0'_{q_w+q_p}]'$ where $0_{q_w+q_p}$ is a $(q_w + q_p) \times 1$ vector of zeros (since we assume that the error terms, ϵ_t, in the PC specified in Equation (2) are independent).

KLS works by considering different reasonable values of r_1, and computing estimates of β that correct for the bias that each value of r_1 entails for standard OLS estimates. In particular, the KLS estimator for β is given by

$$\hat{\beta}_{KLS} = \hat{\beta}_{OLS} - \hat{\sigma}_\epsilon S_{x+x+}^{-1} S_{x+} r,$$

where $\hat{\beta}_{OLS} = (X^{+\prime}X^+)^{-1}X^{+\prime}\pi$ is the OLS estimator of the PC in Equation (2), $\hat{\sigma}_\epsilon^2 = \hat{\sigma}_{\epsilon,OLS}^2/(1 - r'S_{x+}S_{x+x+}^{-1}S_{x+}r)$, $\hat{\sigma}_{\epsilon,OLS}^2 = \frac{1}{T}(y - X^+\hat{\beta}_{OLS})'(y - X^+\hat{\beta}_{OLS})$, $S_{x+x+} = \frac{1}{T}X^{+\prime}X^+$, and S_{x+}^2 is the matrix containing just the main diagonal of S_{x+x+} so that S_{x+} is the positive diagonal matrix $S_{x+}S_{x+} = S_{x+}^2$.

Under the relatively mild regularity conditions in Kiviet (2020) (the strongest one being homoscedasticity), the limiting distribution of the KLS estimator is given by

$$\sqrt{T}(\hat{\beta}_{KLS} - \beta) \xrightarrow{d} \mathcal{N}[0, \sigma_\epsilon^2 V],$$

where V can be consistently estimated by $S_{x+x+}^{-1}\hat{\Theta}S_{x+x+}^{-1}$, and $\hat{\Theta}$ is the sample analogue of the expression for Θ in Kiviet (2020).

To test whether supply-shock proxies have erroneously been excluded or to test whether a candidate IV satisfies the exclusion restriction, we wish to test the hypothesis

$$H_0 : \mathbb{E}[z_{wt}\epsilon_t] = 0 \text{ vs. } H_1 : \mathbb{E}[z_{wt}\epsilon_t] \neq 0, \tag{4}$$

for some $k_w \times 1$ vector of variables z_{wt} (which contains the supply-shock proxies and/or the candidate IVs). Consider the auxiliary regression

$$\pi = X^+\phi_{x+} + Z_w\phi_{z_w} + \nu,$$

where ϕ_{x+} is the new $(1 + q_w + q_p) \times 1$ coefficient vector of X^+, Z_w is the $T \times k_w$ stacked matrix of z_{wt}, ϕ_{z_w} is the corresponding $k_w \times 1$ vector of coefficients, and ν is the $T \times 1$ vector containing the new error terms. Since we are interested in testing whether or not the variables are correlated with the error term, we set $\rho_{[x+z_w]\nu} = [\rho'_{x+\epsilon}, 0'_{k_w}]'$, where 0_{k_w} is a $k_w \times 1$ vector of zeros, run KLS for different values of r_1, and test the exclusion restriction that $\phi_{z_w} = 0$ through a standard Wald test (or a t-test, if $k_w = 1$). As usual, high Wald statistics (lower p-values) imply stronger evidence that the variables in Z_w are in fact correlated with the error term.

Finally, insofar as KLS is simply an endogeneity-corrected OLS estimator, one can use the usual tests developed for OLS to assess whether the assumptions of independent and homoscedastic error terms are fulfilled (see also Kiviet and Kripfganz (2020)). We thus test for heteroscedasticity by regressing the squared residuals at every value of r_1 on the variables included in the PCs, as well as the supply-shock proxies and IVs considered below, and test for autocorrelation by regressing the residuals on different lag lengths of past residuals.

3.4 Data

3.4.1 Main Data Used in the Estimation of the Wage Phillips Curve

The sample we consider runs from 1999Q1 to 2019Q4. Wage inflation is measured as the year-on-year growth rate of total compensation per employee, which is the standard measure used for the analysis of the EA wage PC.[7] For simplicity, we consider the differenced unemployment rate as our measure of slack.[8] Other measures of the slack variable (e.g., the recent Hamilton (2018) filter) are also possible, and do not change the results obtained in this paper substantially. The unemployment rate is the number of unemployed persons as a percentage of the labour force. Price inflation is measured as the year-on-year growth rate of the harmonised consumer price index.

3.4.2 Data on Supply-Shock Proxies and Instrumental Variables

Wage growth is likely to be not only influenced by the development of slack (reflecting mainly changes in demand), but also by supply shocks and monetary policy.

To assess the role of these factors, we include standard variables considered in the literature including oil price inflation (covering global price shocks via energy, see Gordon (2011)), labour productivity growth (reflecting changes in labour supply and its composition, see Nickel et al. (2019)), interest rate spreads (covering changes in the monetary policy stance, see Galì and Gambetti (2019), and exchange rate developments (covering international monetary shocks, see Mazumder (2014)).

In the literature that uses IV methods to test PCs, IVs commonly used are lagged values of the endogenous variable and of the dependent variable, as well as variables such as commodity inflation and the interest rate spread (Mavroeidis, Plagborg-Møller, and Stock, 2014; Kleibergen and

---

[7]As pointed out in Bonam, Haan, and Limbergen (2018, p. 12), negotiated wages rather than compensation per employee are the conceptually better measure, since the latter includes, amongst other things, one-offs, fiscal shocks, and compositional effects. However, contrary to Bonam, Haan, and Limbergen (2018), an EA wage PC (as opposed to country-specific wage PCs) is considered, which makes aggregation of negotiated wages across EA countries more difficult, due to differences in methodology and data quality.

[8]We cannot consider the even simpler measure given by the unemployment rate itself, since for the EA it exhibits non-stationary behaviour, see also Galì (2015).

Mavroeidis, 2009; Dufour, Khalaf, and Kichian, 2006; Galì and Gertler, 1999). In essence, the IV-based approaches in the literature use as IVs precisely those variables that are directly included in the PC by the approaches in the literature that account for the endogeneity of the slack by including supply-shock proxies.

We measure oil price inflation as the year-on-year growth rate of the oil price. Productivity growth is measured as the year-on-year growth rate of productivity per employee, measured as the gross value added in relation to the total number of employees. The interest rate spread is calculated as the difference between the long-term rate and the short-term rate. The exchange rate considered is the USD-euro exchange rate, aggregated to the quarterly frequency by simple averaging. We construct technical controls by interacting these variables as well as their lags amongst each other. We consider two lags of each supply-shock proxy, so that we are left with 78 candidate supply-shock proxies. We construct IVs in a similar way, but following the IV literature on inference on PCs, also include the interaction with the exogenous variables specified in Equation (2) (i.e. the lags of wage inflation and price inflation) and lags of the endogenous slack variable. Throughout, we set $q_w = q_p = 2$. This leaves us with 101 IVs.

These variables are not meant to exhaustively cover all possible supply-shock proxies and IVs. It certainly could be argued that there are other relevant supply-shock proxies and IVs that should be additionally included. However, this illustrates precisely the issue we aim to exemplify.

Indeed, we use this data to show that even with this small and incomplete set of supply-shock proxies, there is substantial specification uncertainty. Arguing that more variables are relevant only worsens the specification uncertainty, and hence strengthens our argument.

Appendix A gives a detailed description of these data, along with their sources.

4   Results

4.1   Single-Equation Estimation Results

Table 1 shows the single-equation estimation results using OLS. Since OLS treats the number of variables as fixed, not all the supply-shock proxies can be considered simultaneously. The only possible approach is to consider the sensitivity of the results to different sets of supply-shock proxies. Column (1) shows the results when no additional supply-shock proxy is added. Columns (2)-(5) show the results when the contemporaneous value, as well as two lags of each respective supply-shock proxy considered are included. Column (6) shows the results when the contemporaneous value of each of the three supply-shock proxies is included (but no lags are included). Column (7) shows the results when five randomly selected supply-shock proxies are chosen from the available 78 technical supply-shock proxies. To keep the results readable, we do not report the point estimates and confidence intervals for the different supply-shock proxies. However, we test the significance of each of the supply-shock proxies included, and report the smallest p-value of these tests in the penultimate row. In the last row we also report the p-value of a Wald test of the hypothesis that the included supply-shock proxies all have a coefficient equal to zero.

The results show that the point estimate of the coefficient of the slack variable is negative across the different specifications considered. However, the 90% confidence sets vary greatly. In column

(6) and column (7) we are unable to reject the hypothesis that the true coefficient of the slack variable is zero. Furthermore, for all cases except for column (4) and column (5), we find that at least one of the included supply-shock proxies has a coefficient significantly different from zero, as can be seen from the the p-values reported in the second row. This suggests two things. First, it suggests that omitted-variable bias caused by omitted supply-shock proxies is a valid concern (although this is not a formal test of the correlation of the supply-shock proxies with the error term of the model without them; see below for a formal direct test of this hypothesis). Second, it suggests that there is no obvious way to select which supply-shock proxies to use. Indeed, there seems to be no principled reason for preferring one set of supply-shock proxies over another, since across the different specifications considered we often find 'significant' supply-shock proxies.[9]

Table 1: OLS single-equation estimation results for the coefficient of the slack variable in Equation (3).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
u	-0.611	-0.619	-0.727	-0.534	-0.863	-0.174	-0.418
	$[-1.061, -0.162]$	$[-1.045, -0.193]$	$[-1.329, -0.125]$	$[-0.904, -0.165]$	$[-1.256, -0.470]$	$[-0.784, 0.436]$	$[-0.748, 0.524]$
t test p-value	$-$	0.002	0.001	0.395	0.269	0.002	0.000
χ^2 test p-value	$-$	0.206	0.085	0.761	0.125	0.552	0.392

Notes:
90% confidence interval based on robust standard errors in brackets.
t test p-value refers to the smallest p-value corresponding to individual significance tests on the coefficients of the supply-shock proxies.
χ^2 test p-value refers to the p-value of a Wald test that tests whether the coefficients of the included supply-shock proxies are jointly equal to zero.
Column (1) does not include any supply-shock proxies. Column (2) uses contemporaneous oil inflation, as well as two lags of oil price inflation, as supply-shock proxies. Column (3) uses contemporaneous labour productivity growth, as well as two lags of labour productivity growth, as supply-shock proxies. Column (4) uses the contemporaneous interest rate spread, as well as two lags of the interest rate spread, as supply-shock proxies. Column (5) uses the contemporaneous USD-euro exchange rate, as well as two lags of the USD-euro exchange rate, as supply-shock proxies. Column (6) uses contemporaneous oil inflation, productivity growth, interest rate spread, and USD-euro exchange rate as supply-shock proxies. Column (7) uses five randomly chosen technical supply-shock proxies. These happen to be the interaction between the one-period and two-period lagged USD-euro exchange rate, the interaction between contemporaneous oil price inflation and the two-period lagged USD-euro exchange rate, the interaction between the one-period lagged interest rate spread and the contemporaneous USD-euro exchange rate, the interaction between two-period lagged oil price inflation and the one-period lagged interest-rate spread, and the interaction between one-period lagged oil price inflation and the two-period lagged interest rate spread.

4.2 Instrumental Variable Estimation Results

Figure 1 shows the plot of one minus the p-value of the AR statistic for different choices of IVs. To avoid overfitting the relationship between the endogenous variable and the IVs, we limit the number of IVs to five.

We present the confidence sets as in Kleibergen (2002). First, we compute the p-value for each hypothesised value of the coefficient of the slack variable in the grid of values considered as outlined in Section 3.2. Second, we report in black one minus the p-value for each of the hypotheses tested. Hence, the 90% confidence set for the coefficient of the slack variable is given by the values of λ_0 for which the black curve is below 0.9, which is shown as the light grey horizontal line. Panel (a) shows the results when the one-period lagged slack variable and the one-period lags of each of the four supply-shock proxies are used as IVs. Panels (b)-(d) show the results when five randomly selected IVs are used.[10]

---

[9]Attempting to 'average' these results across different specifications does not seem feasible. For example, a Bonferonni-type correction would be powerless, since even limiting ourselves to the (few) supply-shock proxies we consider in this paper and to (arbitrarily) setting the number of supply-shock proxies to five, we are left with over 20 million possible specifications.

[10]These happen to be as follows. Panel (b): interaction between one-period lagged wage inflation and the one-period lagged interest rate spread, interaction between one-period lagged wage inflation and one-period
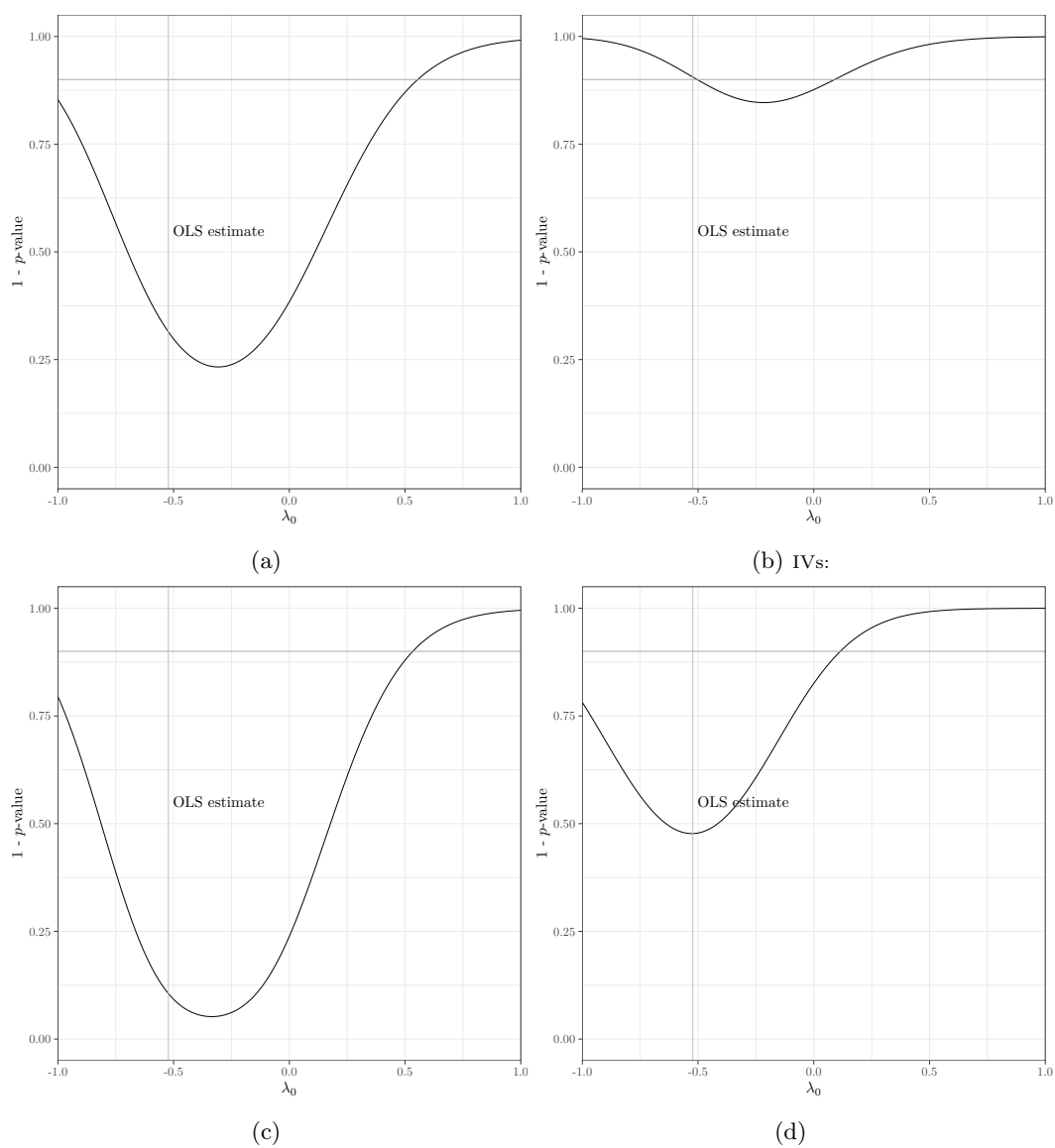
Figure 1: AR confidence sets. See the main text for choice of IVs for each panel.

None of the sets of IVs considered indicate that the slack variable is significantly negative. This suggests a problem of weak identification, which in turn stems from the difficulty of predicting a differenced macroeconomic variable such as the measure of slack we consider. Even so, the results point to substantial heterogeneity in the inferential results. Similarly to the case of single-equation estimation above, this raises the question of how to aggregate the results across the very many combinations possible. Furthermore, although the variables considered above are often used as IVs when conducting inference on PCs, the exclusion restriction is questionable, especially in a reduced-form context.

4.3 Using Kinky Least Squares to Test Phillips Curves Specifications

Figure 2 tests the (joint) hypothesis of uncorrelatedness for the four supply-shock proxies considered in this paper, as well as their lags. The horizontal axis shows the posited correlation coefficient between the slack and the error term, r_1 (see Section 3.3), while the vertical axis shows the p-value of the test of variable exogeneity for the posited degree of endogeneity (see Equation (4) in Section 3.3). Low p-values suggest that the variable considered is correlated with the error term, and that it hence should be included in the wage PC. The results suggest that for large parts of the parameter space of the correlation coefficient between the slack and the error term, the hypothesis that the supply-shock proxies are uncorrelated with the error term can overwhelmingly be rejected. There is no set of supply-shock proxies for which we can conclusively claim that they are not missing from the PC.

Equivalently, this exercise can also be interpreted as a direct test of the validity of the IVs. Figure 3 shows the tests of the individual exclusion restrictions for all of the 101 IVs constructed. The results show that for all degrees of endogeneity, the identifying exclusion restriction of the IVs is unlikely to hold. Therefore, it is not possible to argue for the validity of the IVs by appealing to the intuition that the correlation between the slack and the error term is likely to be positive, since even in this part of the parameter space, the exclusion restriction of these IVs can be rejected.

lagged slack, interaction between one-period lagged slack and one-period lagged productivity growth, interaction between one-period lagged inflation and two-period lagged oil price inflation. Panel (c): interaction between one-period lagged oil price inflation and two-period lagged productivity growth, interaction between the one-period lagged spread and two-period lagged slack, interaction between one-period lagged slack and one-period lagged oil price inflation, interaction between one- and two-period lagged inflation. Panel (d): interaction between one-period lagged slack and one-period lagged oil price inflation, interaction between two-period lagged slack and the two-period lagged USD-euro exchange rate, interaction between two-period lagged wage inflation and two-period lagged inflation, interaction between the one-period lagged USD-euro exchange rate and one-period lagged productivity growth, interaction between one-period lagged oil price inflation and two-period lagged slack.
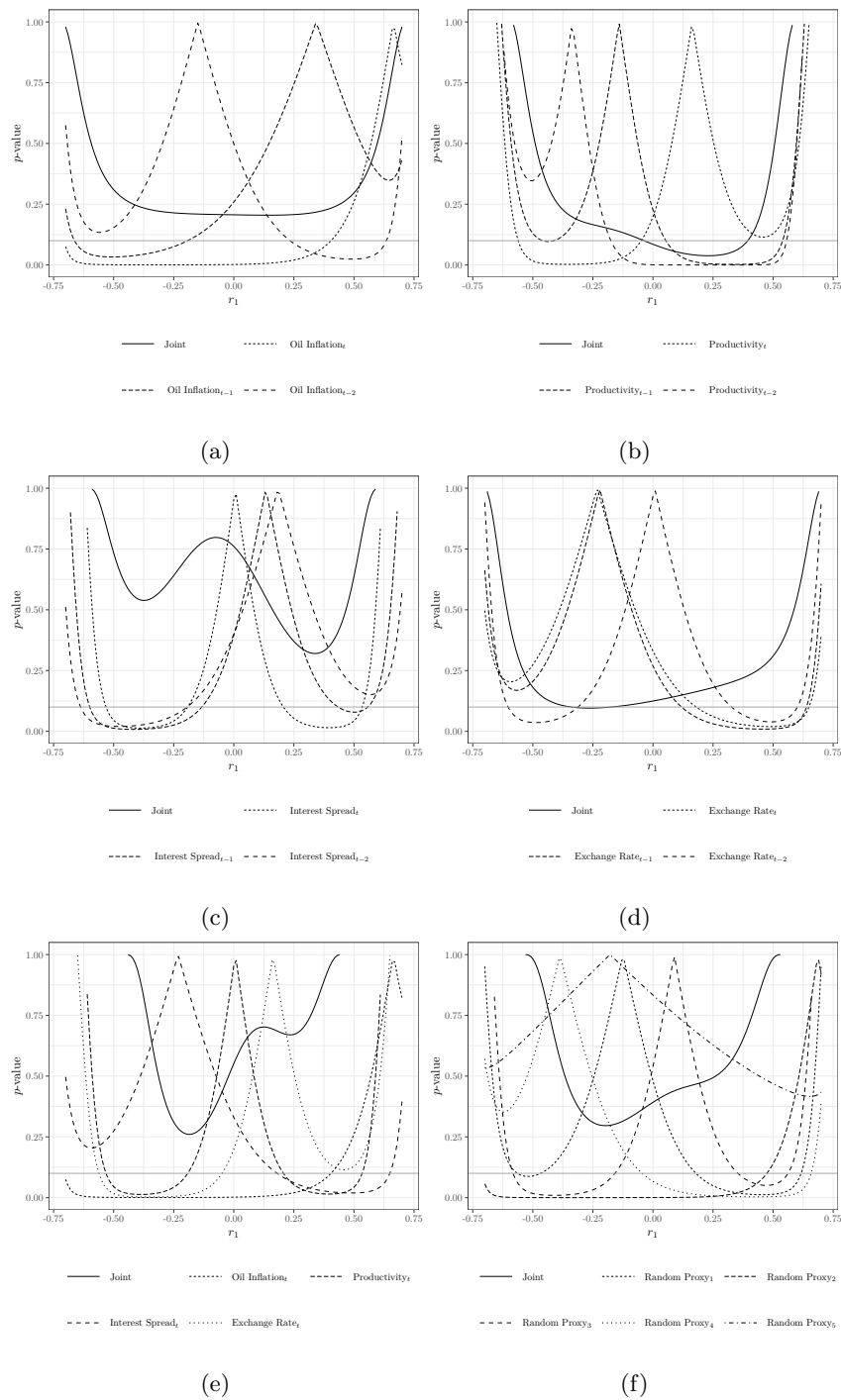
Figure 2: Tests of relevance of supply-shock proxies with KLS. Supply-shock proxies tested in each panel correspond to the ones in Columns (2) - (7) in Table 1.
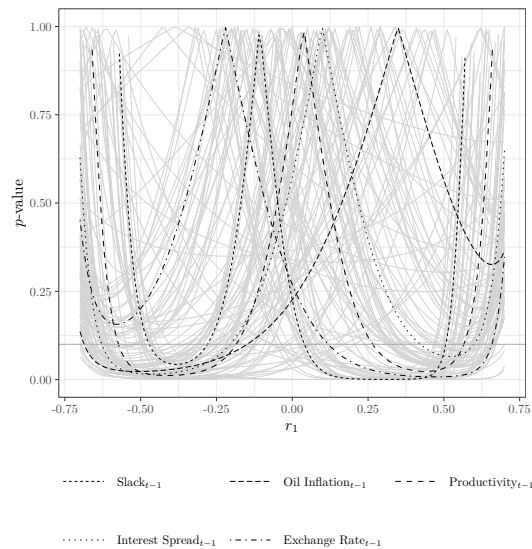
Figure 3: Tests of the exclusion restrictions for the 101 possible IVs. The IVs used for the AR statistic in Figure 1 Panel (a) are shown in black. Other IVs are shown in grey.

Therefore, in line with Gordon (2011) and references therein, we find evidence that supply-shock proxies play an important role in PCs. Furthermore, by directly testing the exclusion restrictions of the IVs, we find evidence that many of the commonly used IVs are in fact not valid. Both these results motivate the KLS estimation in the next section.

4.4 Kinky Least Squares Estimation Results

Figure 4 shows the 90% confidence sets implied by KLS.[11,12] The horizontal axis shows the posited correlation coefficient between the slack variable and the error term, while the vertical axis shows the estimated coefficient of the slack variable (black line) and the 90% confidence interval (area shaded in grey) for the posited degree of endogeneity. Under a 'uniform prior' over the correlation coefficient of the slack variable and the error term, nothing interesting can be said on the likely sign and magnitude of the coefficient on the slack variable. However, if we consider values of r_1 above -0.125, we find that the coefficient of the slack variable is significantly negative.

The slack variable is usually thought to be *positively* correlated with the error term. This can be motivated structurally as in Eser et al. (2020), who show how the standard three-equation New Keynesian model leads to slack being positively correlated with the error term. It can also be motivated by omitted cost-push shocks (Gordon, 2011), a handful of which were found to be missing from our specification in the section above. Under this assumption (or the weaker assumption that the slack is not 'too negatively' correlated with the error term), the results can be interpreted as saying that the coefficient on the slack variable is significantly negative. One obvious point of

---

[11]Considering a broader range does not qualitatively change the results (values of r_1 below or above ± 0.5 lead to point estimates that are distinctly above and below zero, respectively), all while making them less readable.

[12]We note that $r_1 = 0$ corresponds to the OLS estimate and confidence interval.

criticism is that although this assumption seems plausible (and we are unaware of any theoretical or empirical literature claiming that the slack variable should be highly negatively correlated with the error term), in the absence of a valid IV it does remain an unverifiable assumption. However, we nevertheless believe that this approach is more transparent than ones that use a 'preferred specification'. Indeed, while such specifications often rely on ad-hoc arguments that do not have a clear economic rationale (e.g. preferring import prices over exchange rates as supply-shock proxies), our 'identifying assumption' corresponds to a strong prior often held by macroeconomists and policy makers.

Furthermore, inference on the coefficient of the slack variable using KLS offers three advantages over the other inference procedures considered in this paper. First, it addresses all types of endogeneity, and is not limited to addressing the endogeneity caused by omitting the handful of supply-shock proxies considered in this paper. Indeed, although focusing on $r_1 > 0$ is motivated by the important role supply shocks are likely to play, we take no stance on the nature of these shocks, or even require a proxy for them to exist. Second, identification does not rely on questionable exclusion restrictions of the IVs. Third, insofar as no additional variables are used (either as supply-shock proxies or IVs), it also does not suffer from the type of specification uncertainty illustrated above.[13]
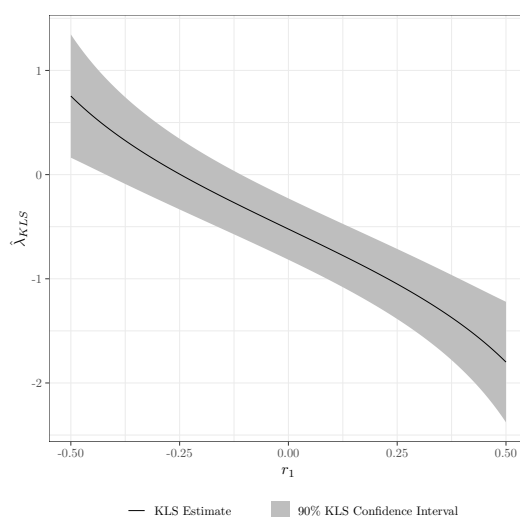

Figure 4: 90% KLS confidence sets for the coefficient of the slack variable in Equation (2).

The validity of KLS depends on the data being homoscedastic and the error-terms being serially uncorrelated. These are both testable assumptions. Figure 5 shows the Breusch-Pagan test for heteroscedasticity as adapted to KLS. The black line shows the p-values that test whether the coefficients of the variables contained in the non-augmented PC (i.e. the variables in Equation (2)) are significantly different from zero when the squared KLS residuals are regressed on them for different values of r_1. The additional variables in transparent grey show the p-values corresponding to the same test, but where one of the 101 variables considered as IVs or the four contemporaneous supply-shock proxies are added (105 variables in total). The p-values are large across all the parameter space considered for all these different combinations, so we conclude that there is no

---

[13]Specification uncertainty stemming from how to measure slack and (wage) inflation of course remains. However, especially in the reduced-form context considered in this paper, this type of specification uncertainty seems unavoidable.

evidence against the assumption of homoscedasticity.

Figure 6 shows the p-values corresponding to the test of whether lagged residuals are significant in a regression of the KLS residuals on lagged values of the KLS residual of increasing order. We find no evidence for autocorrelation for values of r_1 below approximately 0.3, and some evidence for autocorrelation for values between 0.3 and 0.5, and substantial evidence for values above that. Based on recent structural analytic and simulation results in Carvalho, Nechio, and Tristao (2019) and Berriel, Medeiros, and Sena (2016), our informal prior is that values of r_1 above 0.5 can reasonably be discounted. Indeed, assuming that the correlation coefficient between the slack and the error term is above 0.5 almost obviates the need for an empirical analysis of the effect of the slack variable on inflation dynamics: under such a dogmatic prior, standard results on OLS bias in the context of endogeneity would likely be sufficient to reach the same qualitative conclusion that slack negatively affects inflation prior to any empirical analysis.
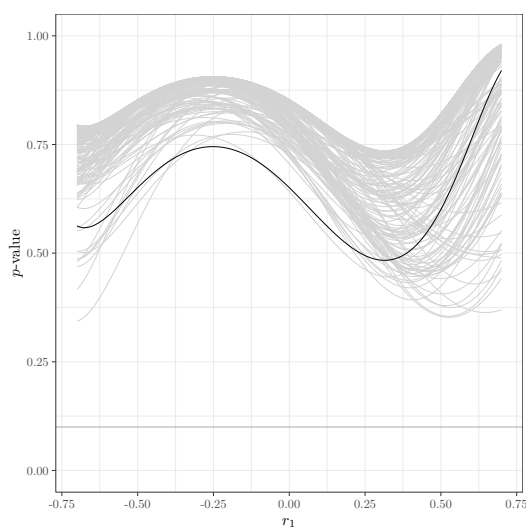
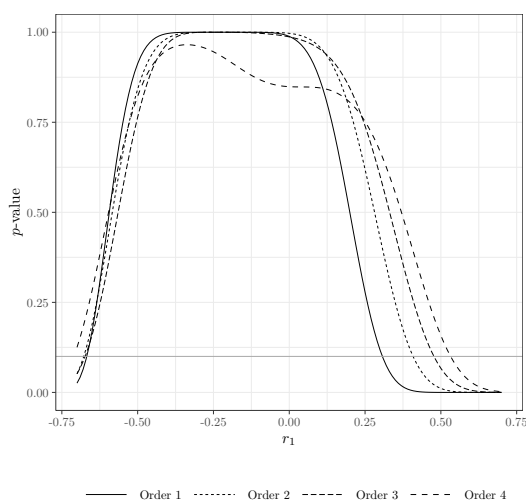Figure 5: Breusch-Pagan test for heteroscedasticity.

Figure 6: Test for autocorrelation.

5 Conclusion

This paper aims to provide endogeneity-robust inference on the coefficient of the slack variable in (reduced-form) PCs. We first identified two broad approaches in the existing literature that seek to account for the endogeneity of the slack variable. The first approach uses proxies of supply-shock variables to address the endogeneity caused by omitting relevant supply shocks, while the second approach seeks to use IVs.

As the first contribution, we showed that conducting inference using such methods is difficult, if not impossible. Approaches that augment standard PCs with supply shocks require the unverifiable assumption that all the endogeneity is due to the variation in the omitted supply-shock proxies, while IV-based methods require making questionable assumptions on the uncorrelatedness of the chosen IVs with the error term. We also showed that both approaches suffer from specification uncertainty, so that it is unclear how to aggregate the inferential results from the very many different specifications available.

The main methodological contribution of this paper to the estimation of PCs consists in advocating the use of KLS. First, we showed how KLS makes it possible to formally test whether certain supply-shock proxies should be included in a PC, and (equivalently) assess whether a variable satisfies the exclusion restriction needed for it to be a valid IV. Second, applying KLS we show that the coefficient of the slack variable in the EA wage PC we analyse is indeed significantly negative, as long as the correlation between the slack variable and the error term is not negative with a large absolute value.

References

Anderson, T and Herman Rubin (Jan. 1949). "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations". In: *The Annals of Mathematical Statistics* 20, pp. 46–63.

Ball, Laurence and Sandeep Mazumder (2019). "A Phillips Curve for the Euro Area". In: *NBER Working Paper* 1.26450, pp. 1–18.

Berriel, Tiago, Marcelo C Medeiros, and Marcelo J Sena (Aug. 2016). "Instrument selection for estimation of a forward-looking Phillips Curve". In: *Economics Letters* 145, pp. 123–125.

Bonam, Dennis, Jakob de Haan, and Duncan van Limbergen (Mar. 2018). "Time-varying wage Phillips curves in the euro area with a new measure for labor market slack". In: *DNB Working Paper*, pp. 1–30.

Bulligan, Guido and Eliana Viviano (July 2017). "Has the wage Phillips curve changed in the euro area?" In: pp. 1–22.

Carvalho, Carlos, Fernanda Nechio, and Tiago Tristao (2019). "Taylor Rule Estimation by OLS". Working Paper.

Chernozhukov, Victor et al. (June 2019). "LASSO-driven inference in time and space". In: *The Annals of Statistics (forthcoming)*, pp. 1–63.

Ciccarelli, Matteo and Chiara Osbat (2017). "Low inflation in the euro area: Causes and consequences". In: *European Central Bank Occasional Paper Series*, pp. 1–89.

Cunningham, Rose, Vikram Rai, and Kristina Hess (Aug. 2019). "Exploring Wage Phillips Curves in Advanced Economies". In: *Bank of Canada Staff Discussion Paper*, pp. 1–27.

Del Negro, Marco et al. (Dec. 2020). "What's Up with the Phillips Curve?" In: *Brookings Papers on Economic Activity*, pp. 301–373.

Dovì, Max-Sebastian (2021). "Inference on the New Keynesian Phillips Curve with Very Many Instrumental Variables". In: *arXiv* 2101.09543, pp. 1–31.

Dufour, Jean-Marie (Jan. 2003). "Identification, Weak Instruments, and Statistical Inference in Econometrics". In: *The Canadian Journal of Economics* 36.3, pp. 767–808.

Dufour, Jean-Marie, Lynda Khalaf, and Maral Kichian (Sept. 2006). "Inflation dynamics and the New Keynesian Phillips Curve: An identification robust econometric analysis". In: *Journal of Economic Dynamics and Control* 30.9-10, pp. 1707–1727.

Eser, Fabian et al. (Sept. 2020). "The Phillips Curve at the ECB". In: *The Manchester School* 88.S1, pp. 50–85.

Galì, Jordi (July 2015). "Hysteresis and the European Unemployment Problem Revisited". In: *NBER Working Paper Series*, pp. 1–51.

Galì, Jordi and Luca Gambetti (Jan. 2019). "Has the U.S. Wage Phillips Curve Flattened? A Semi-Structural Exploration". In: *NBER Working Paper Series*, pp. 1–34.

Galì, Jordi and Mark Gertler (Nov. 1999). "Inflation dynamics: A structural econometric analysis". In: *Journal of Monetary Economics* 44, pp. 195–222.

Gordon, Robert (1982). "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment". In: *NBER Working Paper* 1.708, pp. 1–99.

Gordon, Robert (Jan. 2011). "The History of the Phillips Curve: Consensus and Bifurcation". In: *Economica* 78.309, pp. 10–50.

Hamilton, James D (Dec. 2018). "Why You Should Never Use the Hodrick-Prescott Filter". In: *Review of Economics and Statistics* 100.5, pp. 831–843.

Kiviet, Jan (2020). "Instrument-free inference under confined regressor endogeneity; derivations and applications". In: *Working Paper*.

Kiviet, Jan and Sebastian Kripfganz (2020). "Reassessment of classic case studies in labor economics with new instrument-free methods". In: *Working Paper*.

Kleibergen, Frank (July 2002). "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression". In: *Econometrica* 70, pp. 1781–1803.

Kleibergen, Frank and Sophocles Mavroeidis (July 2009). "Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve". In: *Journal of Business & Economic Statistics* 27.3, pp. 293–311.

Koester, Gerrit et al. (2021). "Understanding low inflation in the euro area from 2013 to 2019: cyclical and structural drivers". In: *European Central Bank Occasional Paper Series*, pp. 1–76.

Ma, Adrian (Aug. 2002). "GMM estimation of the new Phillips curve". In: *Economics Letters* 76, pp. 411–417.

Mavroeidis, Sophocles (May 2005). "Identification Issues in Forward-Looking Models Estimated by GMM, with an Application to the Phillips Curve". In: *Journal of Money, Credit and Banking* 37.3, pp. 421–448.

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James Stock (July 2014). "Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve'". In: *Journal of Economic Literature* 52.1, pp. 124–188.

Mazumder, Sandeep (June 2014). "Inflation in China: Old Versus New Phillips Curves". In: *Europe-Asia Studies* 66.5, pp. 659–709.

McLeay, Michael and Silvana Tenreyro (May 2019). "Optimal Inflation and the Identification of the Phillips Curve". In: *NBER Working Paper Series*, pp. 1–56.

Nickel, Christiane et al. (2019). "Understanding low wage growth in the euro area and European countries". In: *European Central Bank Occasional Paper Series*, pp. 1–88.

A   Data Sources

Table 2: Raw data sources.

Variable	Description	Data Source
Compensation per employee	euro area 19 total compensation of employees	ECB `MNA.Q.N.I8.W2.S1.S1._Z.COM_PS._Z._T._Z.EUR.V.N`
Unemployment rate	euro area 19 unemployment rate (as a % of labour force)	ECB `STS.Q.I8.S.UNEH.RTTT000.4.000`
Consumer price index	euro area HICP overall index	ECB `ICP.M.U2.N.000000.4.INX`
Oil Price	Crude Oil Prices: Brent - Europe (USD)	FRED `DCOILBRENTEU`
Productivity per employee	Gross value added in relation to the total number of employees	ECB `MNA.Q.Y.I8.W2.S1.S1._Z.SAL._Z._T._Z.PS._Z.N`
Short-term rate	euro area Eonia rate	ECB `FM.Q.U2.EUR.4F.MM.EONIA.HSTA`
Long-term rate	euro area 10-year government benchmark bond yield	ECB `FM.Q.U2.EUR.4F.BB.U2_10Y.YLD`
USD-euro exchange rate	U.S. Dollars to one Euro	FRED `DEXUSEU`

Notes:

Interest rate spread calculated as the difference between the long-term rate and the short-term rate.

Inflation rate for variable f_t calculated as $100 \times (\log(f_t) - \log(f_{t-1}))$.

ECB refers to ECB Warehouse Code. FRED refers to FRED Economic Data.

Max-Sebastian Dovì
University of Oxford, Oxford, United Kingdom; European Central Bank, Frankfurt am Main, Germany;
email: max-sebastian.dovi@bnc.ox.ac.uk

Gerrit Koester
European Central Bank, Frankfurt am Main, Germany; email: gerrit.koester@ecb.int

Christiane Nickel
European Central Bank, Frankfurt am Main, Germany; email: christiane.nickel@ecb.int